DESIGN THINKING APPROACH FOR INTERVIEW BOT DEVELOPMENT WITH NATURAL LANGUAGE PROCESSING AND MACHINE LEARNING

DR. S.S.SUGANYA[1], Associate Professor,

LIZORIA HARRIS.A[2], ARAVIND.S[2], AMALORPAVANATHAN. A[2]
Department of Computer Science,
Dr.SNS Rajalakshmi College of Arts and Science (Autonomous), Coimbatore - 49.

Abstract—

The advent of virtual assistants has made communicating with computers a reality. Chatbots are virtual assistant tools designed to simplify the communication between humans and computers. A chatbot will answer your queries and execute a certain computation if required. Chatbots can be developed using Natural Language Processing (NLP) and Deep Learning. Natural Language Process technique like Naïve bayes can be used. Chatbot can be implemented for a fun purpose like chit-chat; these are called Conversational chatbots. Chatbots designed to answer any questions is known as horizontal chatbots and the specific task-oriented chatbots are known as vertical chatbots (also known as Closed Domain Chatbots). In this paper, we will be discussing a task-oriented chatbot to help recruitment team in the technical round of interview process.

Keywords— design thinking, NLP, Machine Learning, Chatbot, python, Google-text-to speech.

I. INTRODUCTION

The use of chat bots has increased extensively in recent years. Many industries, hotels, and flight booking companies use virtual agents to communicate with their users. Chatbots in industries are used for various purposes. Sometimes, they are used to display information. If required, they even help in complex tasks like checking/tracking order status for e-commerce companies. This is the one of the many purposes of task-oriented virtual agents. In this paper, we will be discussing a task-oriented chatbot, which will be useful for companies to filter out the candidates who are less suited for the job based on their scores which will be determined by the bot. Each question will contribute 10% of the total questions (100%). 10 questions in each section: Easy, Moderate and Difficult. We are making use of Naive Bayes as a classification technique. Naive Bayes Classifier uses labels to distinguish between different intents. We will store answers for each question in a label that will be used as intent. Naïve Bayes Classifier is easy and fast. It uses probability to classify text. Naïve Bayes is a probabilistic classifier, which means it predicts on the basis of the probability of an object. This algorithm has a frequency table from which likelihood of the word can be predicted; When salutation is being exchanged, if the user says he is excited for the test, "excited" will be classified as a happy emotion with the help of Bayes theorem. This goal-oriented dialog system will assist the Human-Resources team to eliminate the least suitable candidate as per the company's requirements. A Pattern matching approach is used. In this chatbot to match the answers from the pre-approved answer list which will be stored in the labels. classification method applied in this research is the Naive Bayes method and compared with the Logistic Regression method to determine the class intention.

Leverages the power of Artificial Intelligence and Machine Learning algorithms to create a Speech Recognized Question Answer generator. Besides being a smart and random question generator for test authors, it also serves as a one-stop solution for all content needs of an organization.

II. RELATED STUDY

The following studies have significantly contributed to our understanding of machine learning based technologies such as facial authentication, emotion analysis, sentiment analysis through voice clips, and the development of interview chatbots.

Research by Boucetta et al. (2018) [1] delves into ageinvariant face recognition using deep Convolutional Neural Network (CNN) models, including AlexNet and ResNet50, on the FG-NET database. This study showcases the superior performance of AlexNet, consistently achieving the highest accuracy across age variations, as supported by statistical analysis. Another intriguing facet of facial characteristics is explored in Oosterhoff et al.'s (2008) [2] research, which investigates the relationship between facial features and social evaluations. It reveals that specific facial attributes, such as the curvature of the mouth, can influence how individuals perceive trustworthiness and dominance, with contextual connotations that can be either positive or negative.

Transitioning to the realm of interview chatbots, S. G. and Y. Chen's study (2022) [3] examines the use of chatbots in facilitating inclusive learning through an interview study. This work underscores the advantages of employing Natural Language Processing (NLP) to develop an interview bot capable of engaging with candidates. The bot's workflow involves generating main and sub-questions and recording candidate responses through the keyboard. In parallel, J. Siswanto et al. (2022) [4] investigate the development of an interview bot employing NLP and machine learning techniques. Their research delves into the effectiveness of chatbots equipped with active listening skills and assesses the feasibility and efficacy of building interview chatbots using publicly available technologies. This study concludes that integrating active listening skills into chatbots enhances user engagement and leads to higher-quality responses.

Furthermore, M. X. Z. C. Y. C. Ziang Xiao's work (2020) [5]emphasizes the enhancement of interview chatbots through active listening skills. The research highlights that incorporating active listening capabilities significantly improves user engagement and results in higher-quality user responses.

Recent advancements have focused on leveraging machine learning strategies to augment chatbots' natural language processing capabilities. Suakanto et al. (2021) [6] present a novel reinforcement learning-based approach for training chatbots. This method involves instructing chatbots on effective communication within a simulated environment, enabling them to adapt to various user requests. The authors suggest that this approach equips chatbots to handle complex requests more effectively and adapt their responses as needed.

III. System Methodologies

The research follows an experimental design to develop and evaluate the Interview Bot to revolutionize the hiring process and facilitate a more efficient and interactive interview experience for candidates. Instead of traditional and existing text-based questionnaire type interviews, the bot presents questions on the screen, and candidates can respond using speech, simulating a real-life interview scenario. The research methodology of the Interview Bot encompasses four main components, speech recognition, facial recognition, question categorization, and the grading system. Each component plays a crucial role in achieving the system's objectives and improving the efficiency and accuracy of the hiring process.

A. System Overview

This web-based application as illustrated in figure [1], provides a user-friendly interface for candidates to access interviews using their designated company credentials. Once logged in, candidates engage in an interactive interview process, responding to a series of questions posed by the system. A distinct highlight of this process is the system's ability to record video responses, capturing not only verbal answers but also facial expressions and vocal nuances.

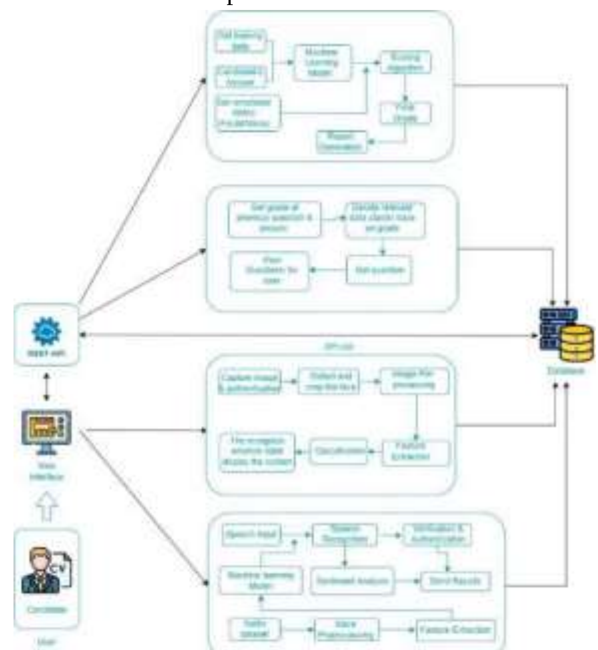


Figure 1: System Overview Diagram

Through intricate machine learning algorithms, it conducts a comprehensive evaluation, encompassing sentiment, emotion, and voice analysis. Furthermore, the system diligently scans the video for any signs of cheating, ensuring the integrity of the interview process. The culmination of this analysis generates a detailed report, offering nuanced insights into the candidate's performance. This comprehensive report is then sent to the company via email. With this report in hand, companies gain a better understanding of each candidate's suitability.

B. Data Collection and Preprocessing

The data collection process for the voice and facial components of the Interview Bot involved acquiring carefully selected datasets from Kaggle.com, a reputable platform renowned for hosting a diverse array of publicly available datasets. For the voice component, the dataset encompasses audio recordings exhibiting diverse sound characteristics and a range of emotions/sentiments, categorized primarily into four types as calm, happy, discontent, and neutral. On the other hand, the facial component dataset encompasses many facial images, capturing a range of expressions and emotions effectively as positive, negative, and neutral. For the question bank and categorization, the data collection process involves a combination of manual curation from wide range of publicly available resources. A diverse set of interview questions is compiled, covering various technical, behavioral, and situational aspects. These questions are manually chosen by experts in the domain, to ensure relevance and effectiveness in assessing candidates' skills.

C. Voice Recognition and Sentiment Analysis

To enable speech-based interaction with the Interview Bot, we incorporated speech recognition technology and algorithms. The methodology includes feature extraction, model creation, data loading and preprocessing technologies. Feature extraction involves using the librosa library to extract essential characteristics from each audio file in the dataset, such as Mel Frequency Cepstral Coefficients (MFCC), Chroma, and Spectral Contrast features, which are commonly used in audio/speech analysis. The extract feature's function is utilized to load the audio file and extract these features. Model creation focuses on defining an LSTM model that consists of two LSTM layers followed by a Time Distributed Dense layer, a Flatten layer, a Dropout layer for regularization, and a final Dense layer for classification. This model architecture enables the capture of temporal dependencies in audio/speech data, which is crucial for accurate analysis. In the data loading and preprocessing step, the dataset is organized with one subdirectory per emotion, and audio files corresponding to each emotion are processed to extract features and create a feature matrix (X). The target labels (y) are then one-hot encoded to facilitte classification. The next step involves answer analysis, where the system employs advanced Natural Language Processing (NLP) algorithms to analyze candidates' responses and send that data to the grading system to compare the accuracy of the given answer with correct answer in the system. These steps collectively enable effective speech recognition and sentiment analysis within the interview bot system as depicted in the flowchart shown in Figure 2.

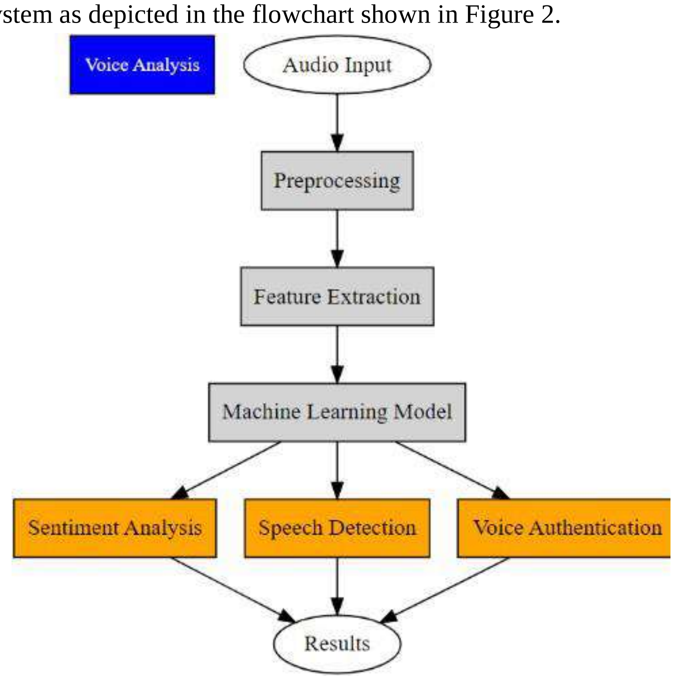


Figure 2: Voice Analysis Overview

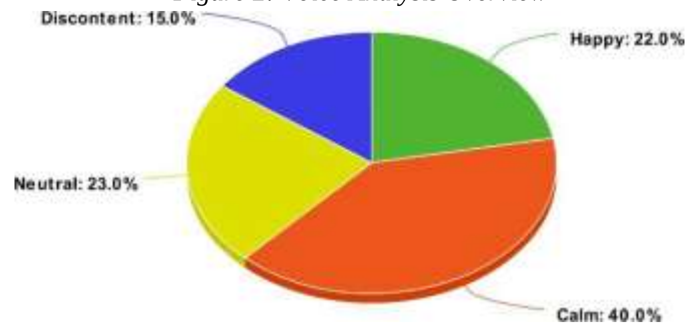


Figure 3: Sentiment Analysis Test Results for an Interview

D. Facial Recognition Analysis

The facial recognition component of the system was developed using a systematic methodology. The first step involved sourcing a facial recognition dataset from a reputable platform, with Kaggle being the chosen platform for this purpose. This dataset consisted of diverse facial images with labeled emotion categories, ensuring a comprehensive representation of different expressions. A pre-trained Keras model, specifically trained to recognize three emotion classes (Negative, Neutral, and Positive), was loaded. Facial images were loaded using OpenCV and converted to grayscale to match the expected input format of the model. The Haar Cascade frontal face detector was utilized to detect faces in the images, providing bounding box coordinates. For each detected face, the image was cropped, resized to a standardized size, converted to a numpy array, and normalized. The preprocessed face images were then fed into the pre-trained model to obtain predictions for each emotion class. The prediction with the highest probability was selected, and the corresponding emotion label was extracted and printed as the predicted emotion for the detected face.

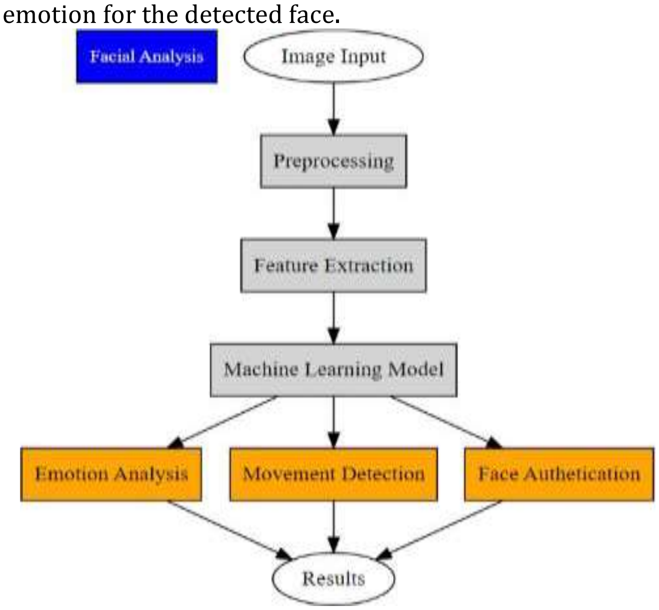


Figure 4: Facial Analysis Overview

E. *Grading System Analysis*

The grading system component of the research paper incorporates a methodology that aims to achieve an objective and consistent evaluation of candidates' responses during the interview process. The methodology begins with the establishment of criteria, which cover important aspects such as problem-solving ability and technical expertise. These criteria are thoughtfully selected to encompass the essential qualities required for the job position. By identifying and extracting relevant keywords and phrases from the data received by the voice recognition component the system objectively assesses candidates' performance against the predefined criteria. The grading system employs a machine learning model that has been trained on a comprehensive dataset comprising past interviews. This model considers the extracted keywords, phrases, and other pertinent features from candidate responses to calculate scores for each criterion. To provide a holistic evaluation, the grading system consolidates the scores obtained from individual question evaluations. Weighting factors may be applied to different criteria based on their relative importance as shown in Table [1] below. The grading system incorporates the evaluation of emotional and sentiment analysis scores obtained from both facial and voice analysis. These scores provide additional dimensions for assessing candidates' performance, considering their emotional engagement, communication style, and overall sentiment expressed during the interview.

Numerical Grade	*Grade*	*Next Question Category*
75 < X < 100	A	Difficult
50 < X < 75	B	Medium
0 < X < 50	C	Easy

Table 1: Grading System

F. *Question Bank Analysis*

The development of the question bank and categorization process involves a systematic approach aimed at enhancing the interview bot's efficacy and adaptability. A diverse set of interview questions relevant to various job positions in the IT sector is collected. These questions are obtained through collaboration with industry experts, HR professionals, and domain-specific resources. The questions cover technical skills, behavioral traits, problem-solving abilities, and other relevant aspects. This raw question pool forms the basis for the question bank.

To initiate this component, data loading and structuring are performed using the Pandas library. The 'questions_answers.csv' file, containing pairs of interview questions and their respective answers, is loaded into a Pandas DataFrame. This structured data format allows for systematic organization and manipulation of the interview content. This serialized form of data storage enhances efficiency in data retrieval and management, enabling seamless integration with the broader system.

A pivotal aspect of the methodology lies in the categorization of interview questions. The 'level' column in the DataFrame serves as a criterion for categorization based on question complexity. By classifying questions into distinct complexity levels such as 'simple,' 'medium,' and 'complex,' the system can tailor interview experiences to candidate skill levels, optimizing the assessment process.

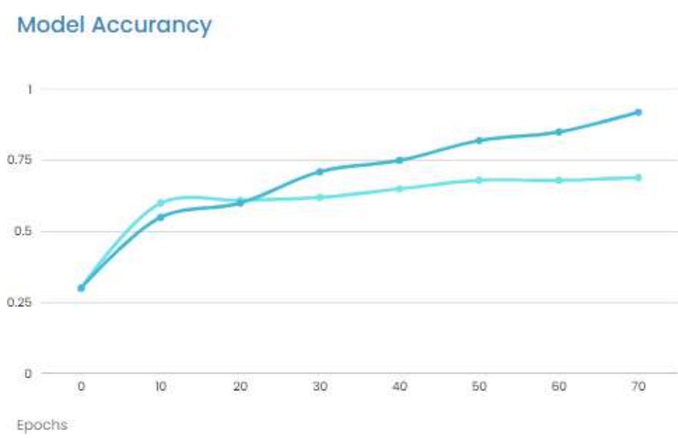
IV. DESIGN THINKING

RESULTS & IMPLEMETATION

The study's findings shed light on the system's performance across its key components which facilitate efficient and standardized candidate assessments.

A. *Speech Recognition Component Results*

The speech recognition component demonstrated promising accuracy in converting spoken language to text. In a sample test with diverse voice recordings, the system accurately transcribed speech with an average accuracy of over 60%. Sentiment analysis performed on the transcribed text, involving pitch and voice features analysis, exhibited an accuracy rate of approximately 70%. This suggests the system's capability to determine the emotional tone of candidate responses.



V. CONCLUSION AND FUTURE WORK

Although traditional human interviewing method is accepted worldwide, it can be replaced to some extent by using AI powered chatbot. System is developed to use in IT companies as of now. Companies would definitely select this system to overcome the drawbacks in existing systems and procedures. Furthermore system will bring transparency and in candidate selection process. Chatbot will generate results within minutes as compared to traditional method. As knowledge-base is connected to internet, it can be easily updated and changes would take place immediately.

Advancements in NLP and ML are propelling chatbots toward a future where their conversations closely mimic human interaction. Users will experience more natural and engaging interactions with chatbots, improving user satisfaction.

Because of the scope of the project we did not have time to conduct as much user testing and re-design to the chatbot as we would have liked. This has an impact on the validity of our research. Through the project we have touched on some theory when making the chatbot, but this should also have a larger focus for higher validity. Even though the participants trusted the information given in this project we cannot say that people trusts a chatbot as much as they trust a human being. There are also biases in our project, one of them is that all the students that we included in the project already knew a lot of the answer the prototype could provide. Another bias is that the information the chatbot provides could be seen as "casual" and are not crucial and/or vital This could have had an impact on the results regarding trustworthiness. With that being said we also think that some of our findings could give some insights into how a very small group of people think about using a chatbot to gain information in a school context. Some of the characteristics of our chatbot was viewed as appropriate for the given context, like "casualness" and links to where the information was gathered. If the IFI chatbot is to be furthered developed, this could be something to draw upon.

VI. FUTURE WORKS AND SUGGESTIONS

While our study has notably concentrated on the software engineering domain, the horizon for future advancements is expansive. A promising direction for enhancement lies in diversifying the interview domain. By extending the system's training to encompass a broader spectrum of professions and industries, we could unlock its potential to serve a more diverse array of recruitment needs. Such an evolution would necessitate an augmented question bank that caters to various domains, thereby widening the application landscape and rendering the Interview Bot adaptable to a wider array of industries. Another pivotal avenue is the integration of multi-language support, leveraging advanced natural language processing techniques to facilitate interviews conducted in diverse languages. This expansion would amplify the system's accessibility on a global scale, enabling organizations worldwide to harness the capabilities of the Interview Bot seamlessly. Empowering organizations to tailor the evaluation criteria according to their specific requirements would enhance the system's adaptability and suitability across various domains.

REFERENCES

1) Joko Siswanto, Sinung Suakanto, Made Andriani IJtech Vol 13, No 2 (2022), "Interview Bot Development with Natural Language Processing and Machine Learning",

2) Suman K. Saksamudre, P.P. Shrishrimal, R.R. Deshmukh, "A Review on Different Approaches for Speech Recognition System"

3) S. Furui; T. Kikuchi; Y. Shinnaka; C. Hori 2004 - ieeexplore.ieee.org , "Speech-to-text and speech-to-speech summarization of spontaneous speech".

4) Augello, A., Gentile, M., Weideveld, L., Dignum, F.: A model of a social chatbot. In: De Pietro, G., Gallo, L., Howlett, R.J.,Jain, L.C.(eds.) Intelligent Interactive Multimedia Systems and Services 2016. SIST, vol. 55, pp. 637–647. Springer, Cham (2016)

5) [Nishimura et al., 2005] R. Nishimura, Y. Watanabe and Y. Okada. A Question Answer System Based on Confirmed Knowledge Developed by Using Mails Posted to a Mailing List. In Proceedings of the IJCNLP 2005,pp.31-36,2005

6) Avalverde, D. (2019). A Brief History of Chatbots. Perception, Control, Cognition. Retrieved March 9, 2019 from: https://pcc.cs.byu.edu/2018/03/26/a-brief-history-of-chatbots/

7) Ayedoun, E., Hayashi, Y., & Seta, K. (2015). A Conversational Agent to Encourage Willingness to Communicate in the Context of English as a Foreign Language. Procedia Computer Science, 60(1): 1433–1442.

8) Ben Mimoun, Mohammed Slim, & Poncin, I. (2015). A valued agent: How ECAs affect website customers' satisfaction and behaviors. Journal of Retailing and Consumer Services, 26: 70– 82.

9) Chatbot Magazine (2019). A Visual History of Chatbots. Retrieved March 9, 2019 from: https://chatbotsmagazine.com/a-visual-history-of-chatbots-8bf3b31dbfb2

10) Colace, F., De Santo, M., Lombardi, M., Pascale, L., Pietrosanto, A. (2018). Chatbot for E-Learning: A Cases Study. International Journal of Mechanical Engineering and Robotics Research Vol. 7, No. 5, September.

11) Egencia (2018). What is a Chatbot and How does it work? Retrieved March 9, 2019 from: https://www.youtube.com/watch?v=38sL6pADCog

12) Hattie, J. (2012). Visible learning for teachers: Maximizing impact on learning: Routledge. https://chatbotsmagazine.com/a-visual-history-of-chatbots-8bf3b31dbfb2

13) Lip ko, H. (2018). Meet Jill Watson: Georgia Tech's first AI teaching assistant. Retrieved on March 9, 2019 from: https://pe.gatech.edu/blog/meet-jill-watson-georgia-techs-first-ai-teaching-assistant.